Influence of Information Structure on the Salience of Opinions

Josef Ruppenhofer, Daniela Schneevogt
Marienburger Platz 22
Dept. of Information Science and
Natural Language Processing
Hildesheim University
31141 Hildesheim
ruppenho@uni-hildesheim.de, d.schneevogt@gmx.de

Abstract

We study the influence of information structure on the salience of subjective expressions for human readers. Using an online survey tool, we conducted an experiment in which we asked users to rate main and relative clauses that contained either a single positive or negative or a neutral adjective. The statistical analysis of the data shows that subjective expressions are more prominent in main clauses where they are asserted than in relative clauses where they are presupposed. A corpus study suggests that speakers are sensitive to this differential salience in their production of subjective expressions.

1 Introduction

It is well known that subjectivity or sentiment is a complex phenomenon. Not only do individual subjective expressions such as *handsome, beautiful, ugly* differ in intensity and polarity, but also external factors impinge on subjective expressions, modulating their intensity and/or polarity. Accordingly, the best studied questions in this area of research include methods for assigning prior polarity (Hatzivassiloglou and McKeown, 1997) and for recognizing polarity in context (Wilson et al., 2005; Moilanen and Pulman, 2007) ; methods for assigning out-of-context (or: prior) polar intensity scores to adjectives (Sheinman and Tokunaga, 2009; de Melo and Bansal, 2013; Ruppenhofer et al., 2014) and methods for modeling the effects of degree modification (Taboada et al., 2011). Negation and modality have been studied by (Benamara et al., 2012; Wiegand et al., 2010).

Greene and Resnik (2009) look at the influence of what they call syntactic framing on subjectivity, namely questions of causal responsibility, affectedness and salience of new states resulting from events.

What has, to our knowledge, not been investigated at all is the influence of `information structure` on the salience of subjective expressions.[1] Information structure (Lambrecht, 1996) is that part of linguistics that concerns itself with the relation of sentence form to the linguistic and extra-linguistic contexts in which sentences are used to convey propositional information. Importantly, sentences may contain the same propositional information, yet differ in terms of information structure, as shown by examples (1-2).

(1) Peter, who is a really **sweet** guy, lives next door.

(2) Peter, who lives next door, is a really **sweet** guy.

Both sentences convey to the hearer new information about a known topic, namely the referent of *Peter*. Importantly, in each sentence the information in the relative clause is presupposed, that is, presented as if it already is part of the so-called common ground between speaker and hearer. By contrast, the main clause predicate is part of the focus, i.e. it is assumed to provide new information about the topic. The question that we investigate experimentally in this study is whether subjective expressions that are asserted as part of the focus are perceived more saliently than subjective expressions that are embedded in presupposed parts of sentences. In other words, are speakers likely to rate (2) as more subjective than (1) due to *sweet* being part of the focus in the former but not the latter?

Our experiments show that there are clear and stable differences between sentences that contain all the same lexical material and all the same propositions but which structure these propositions in different ways.

This paper is structured as follows. Section 2 presents related work. In section 3, we lay out our experimental design. The results of our data collection are analyzed and discussed in section 4. A supplementary corpus study is presented in section 5 and we conclude in section 6.

2 Related Work

One type of related work looks at how evaluative text is structured and where subjective expressions may be found. Bieler et al. (2007) develop a system for analyzing movie reviews into formal (e.g. author, genre, legal-note etc.) and functional parts (describe vs comment). Degaetano-Ortlieb et al. (2012) study the type and distribution of sentiment expressions occurring in different sections of texts. Their purpose is to study the similarities and differences between related but different scientific disciplines.

Wang et al. (2012) seek to incorporate information about discourse relations such as Contrast, Cause, etc. into the task of classifying reviews. The knowledge they use is pragmatic in nature, but it is orthogonal to information structure. Similarly, Mukherjee and Bhattacharyya (2012) use discourse relation information gleaned without full discourse parsing to improve tweet polarity classification.

Heerschop et al. (2011) use Rhetorical Structure Theory to divide a text into important and less important text spans, subsequently using this to improve the performance of a sentiment classifier. The discourse relations of RST concern the propositional content of pairs of propositions. By contrast, information structural notions such as topic, focus, presupposition and assertion are properties of individual clauses as they concern the relation of a proposition to the knowledges state of the speaker and hearer about the content of that proposition. The effects of information structure are thus distinct from the effects of the kind of discourse structures that RST covers.

Wiebe and Riloff (2005) classify sentences as subjective and objective based on extraction patterns they learn. This work operates on the sentence level but it looks at the detection of subjectivity rather than its salience. It uses extraction patterns but does not model where they occur within the sentences.

Our work can also be compared to work on determining the intensity of subjective expressions, some of which we referenced in the introduction. We work on contextual effects. However, they are effects on sentence subjectivity rather than the intensity of subjective expressions, and we study information structure as an influence rather than the effect of degree modification or negation.

Finally, Kabadjov et al. (2011) investigate the suitability of very highly positive and negative sentences for the purposes of text summarization. They find that sentences found useful for summarization are no different in terms of subjectivity intensity than sentences that were not found similarly useful. While this work looks for salient sentences that are useful for the summaries, it does not take into account how prominent subjective expressions are within the sentences that are either salient for summary purposes or not. Thus, the issue that interests us is not addressed.

3 Experimental design

3.1 Selecting the clause types

The purpose of the study is to test the influence of information structure on the salience of subjective expressions. Information structure can be signaled through various linguistic means, including intonational and lexical means. Since we were looking to perform a self-paced reading experiment and wanted to avoid possible confounding influences introduced by lexical cues to information structure, we decided to focus on different sentence types as signals of particular information structures.

We specifically contrast main clauses and relative clauses. We chose these two clause types because they were among the most common types in a 250 sentence random sample taken from the Huge German Corpus (HGC; Fitschen 2004), which contains around 204M words of newspaper text. The focus on these clause types is meant to avoid any distortion of the results through low-frequency structures. Note that we work only with asserted indicative mood sentences so as to exclude modality as a variable influencing our results.

We did not use complement clauses in our study, although they were among the most frequent clause types in our HGC sample. The reasons for this choice are the following. First, constructing complement clause stimuli would mean using different/additional lexical material, whereas main and relative clause pairs can be constructed so they contain the same lexical material (cf. 3–6). Second, since complement clauses vary by the type of embedding predicate – e.g. whether it is factive (e.g. *know*) or not (e.g. *claim*) – one would need to control for the various subtypes of complement clauses, which would increase the amount of items to be rated and thus the length of the survey. Third, complement clauses with predicates of cognition or communication present the additional difficulty that the source of the subjective expression is an attributed one rather than the sentence's author, whereas for the relative and main clause data the source is always the implicit author. For this first study of information structural influences on sentence subjectivity, it was thus easier, both in terms of stimuli construction and subsequent analysis, not to include complement clauses.

3.2 Selecting the adjectives

Adjectives are a well-studied lexical class clearly associated with evaluation and subjectivity (e.g. (Bruce and Wiebe, 1999; Wiebe, 2000)). They are often the largest class in polarity lexicons (e.g. SoCAL (Taboada et al., 2011) or the Pittsburgh subjectivity clues (Wilson et al., 2005)), and opinion mining systems that limit themselves to use only words of certain parts of speech as features will tend to include adjectives. Accordingly, we decided to focus our experiments on the salience of adjectives in various configurations.

We created a pool of candidate adjectives by merging the adjectives contained in two German polarity dictionaries, SentiWS (Remus et al., 2010) and German Political Clues (Waltinger, 2010). Since we wanted to control for the inherent polarity strength of the adjectives, we decided to select the adjectives in pairs such that both refer to the same semantic scale but one has greater prior intensity than the other. For example, *spannend* 'fascinating' is more positive than *interessant* 'interesting' but both refer to the scale of mental stimulation. Further, in order to control for the influence of word familiarity on the results, we evaluated the frequency of our candidate adjective pairs in two ways. First, we checked against the dlexDB psycholinguistic database (Heister et al., 2011) to make sure they had about the same frequency of occurrence there. And second, we checked that both adjectives were within the same frequency band for adjectives in the HGC corpus. Finally, to avoid results due to lexical idiosyncrasies, we constructed 5 positive and 5 negative pairs of adjectives. We also chose 8 different control adjectives that we expected to be neutral given that they were not listed in the two German polarity lexicons we used.[2] The chosen adjectives are listed in Table 1. The polar adjective pairs are indicated by way of horizontal lines.

3.3 Stimuli

We illustrate our experimental stimuli with the set of examples in (3)–(7). The codes preceding the examples are constructed as two-letter combinations as follows: R = relative clause, M = main clause, C = complement clause; W = weak prior intensity, S = strong prior intensity; N = neutral. In the interest of keeping the length of our survey at around 20 minutes, we decided not to use RN and MN stimuli, since we are most interested in the behavior of the non-neutral adjectives across conditions. We included the neutral sentences as filler material. Note that sentences with polar adjectives were constructed so they do not contain any other subjective expressions.

(3) [RS] Ihr Bruder, der zu allen **unfreundlich**

[2] This expectation was not fully borne out in the experiments, as will be discussed in section 4.3.

Adjective	Gloss	Polarity SentiWS	Polarity elicited
ungeschickt	clumsy	-0.6087	-34.2576
doof	daft	-0.1562	-50.8333
unfreundlich	unfriendly	-0.3407	-62.3485
unhöflich	impolite	-0.0048	-60.2727
eintönig	humdrum	-0.0378	-42.6212
langweilig	boring	-0.0228	-49.6970
entsetzlich	appalling	-0.477	-71.1509
scheußlich	hideous	-0.1834	-75.8868
dumm	dumb	-0.5901	-61.6038
blöd	stupid	-0.1593	-55.4528
hübsch	pretty	0.4629	56.8636
wunderschön	gorgeous	0.7048	78.0152
grandios	grand	0.1843	80.1515
großartig	great	0.4606	78.7879
freundlich	friendly	0.6022	65.2830
nett	nice	0.1405	49.2642
intelligent	intelligent	0.1238	65.6038
klug	smart	0.3532	64.5094
interessant	interesting	0.2488	51.0377
spannend	fascinating	0.7165	50.6415
geheim	secret	neutral	-2.3940
geläufig	prevalent	neutral	4.2576
wissenschaftlich	scientific	neutral	22.4848
gängig	common	neutral	5.3788
objektiv	objective	neutral	19.4340
sachlich	matter-of-fact	neutral	14.8491
häufig	frequent	neutral	2.4340
dünn	thin	neutral	2.2453

Table 1: Adjectives used in human elicitation

ist, wohnt in Berlin. ‘Her brother, who is unfriendly to everybody, lives in Berlin.’

(4) [RW] Ihr Bruder, der zu allen **unhöflich** ist, wohnt in Berlin. ‘Her brother, who is impolite to everybody, lives in Berlin.’

(5) [MS] Ihr Bruder, der in Berlin wohnt, ist zu allen **unfreundlich**. ‘Her brother, who lives in Berlin, is unfriendly to everybody.’

(6) [MW] Ihr Bruder, der in Berlin wohnt, ist zu allen **unhöflich**. ‘Her brother, who lives in Berlin, is impolite to everybody.’

(7) [CN] Sie erzählt, dass ihr Bruder in Berlin wohnt. ‘She says that her brother lives in Berlin’.

3.4 Task

Our contained main task and distractor task items. The *main task* consisted in rating sentences on a 7 point scale ranging from strongly negative (-3) via neutral ∅ to strongly positive (+3). The survey was administered to the subjects via the open-source LimeSurvey[3] software. The sentences were organized in groups (such as (3)–(7)) and we randomized both the ordering of the sentences within a given group as well as the ordering of the groups within the survey. This was done to control biases that might arise due to learning, habituation or motivational effects in the course of answering the survey questions. The sentences were displayed singly on the screen and subjects had to press a continue-button to go on to the next item. They could not return to a previous item. Though displaying multiple items per screen would have made completion of the survey faster, grouped display is likely to result in respondents viewing the items as a set and increasing the correlation among them beyond what is due to the stimuli themselves (cf. discussion by Couper et al. (2001)).

In the *distractor task*, subjects were asked to rate the polarity and intensity of adjectives that appeared in the main task. The scale for the adjective intensity rating ranged from -100 to +100. The purpose of the distractor task was to a) prevent participants from focusing consciously too much on the sentence rating task; b) check on the information available from the polarity lexica ; and c) use values for prior adjective intensity that actually fit our population of subjects.

Since we did not have funds to pay our subjects, we tried to keep the duration of the survey within a 20-25 minute time window. Because a pre-test had shown that our initial design took longer than 20 minutes for many participants, we divided our main and distractor task items across two non-overlapping versions A and B for the actual run of the survey. Version A included 3 positive adjective pairs and 2 negative pairs, Version B covered 2 positive adjective pairs and 3 negative ones. Using two complementary surveys allowed us to elicit ratings for the adjectives in the sentence rating task of survey A in the distractor task of survey B, and vice versa.

Besides keeping the survey short in the interest of a higher completion rate, subjects also were shown a progress bar on the screen so they would not abandon surveys when they were already close to completion. In addition, some extrinsic motivation to complete the survey was provided by the chance for participants to win one of three vouchers for use at a large online merchant.

3.5 Subjects

Our subjects mainly are undergraduate students at two German universities. These participants were recruited from the friends and acquaintances of the au-

[3]Available at www.limesurvey.com/.

thors and also via colleagues and referrals from participants. We collected meta-data (gender, age, German proficiency, place of birth and place of residence, occupation) on our subjects but do not use them in our analysis below. Overall, 130 subjects completed the survey altogether. 72 completed survey A and 58 survey B. Note that subjects were assigned randomly to the two versions of the survey.

4 Results and discussion

4.1 Data cleanup

In order to eliminate the influence of participants who might have had difficulty with the task, we proceeded as follows. For each participant we calculated the average of their kappa values with every other subject. Based on the mean and standard deviation of these average kappa values, we excluded those participants that lay outside 2 SDs of the average kappa of the average annotator. In addition, to err on the side of caution, we also decided to exclude the ratings of participants for whom German is not their native language. As a result, we retained the data of 110 participants, 64 from version A and 46 from version B.

4.2 Analysis

Figure 1 shows two boxplots for the absolute adjective ratings grouped by sentence type, with the results for non-polar sentences on the left and for polar ones on the right.[4] These results seem in line with our expectations: for non-polar adjectives there are no significant differences. For polar ones, there are differences among the main clauses and relative clauses.

Figure 2 shows plots of average word intensity ratings against average absolute sentence ratings with one dot representing each adjective. The plots show that, as expected, higher adjective intensity goes with greater sentence subjectivity in both relative and main clauses. The comparison of the slopes in the graphs also suggests that the relationship is somewhat stronger in main clauses than in relative clauses.

Since each adjective has a unique absolute prior intensity and each adjective appears with both sentence types, we can draw an imaginary vertical line through an adjective's intensity value on the x-axis and see how the adjective's average sentence rating in main clauses compares to its rating in relative claues. Doing this shows that the average sentence ratings in the main clauses are greater than the ratings in the relative clauses for 19 of the 20 polar adjectives. The remaining case is the adjective *ungeschickt* 'clumsy' in the lower left corner of the graph. For this item, which was rated as the least intense adjective among the ones that we had taken to be polar based on the polarity

[4]The black dots in the plots represent outliers.

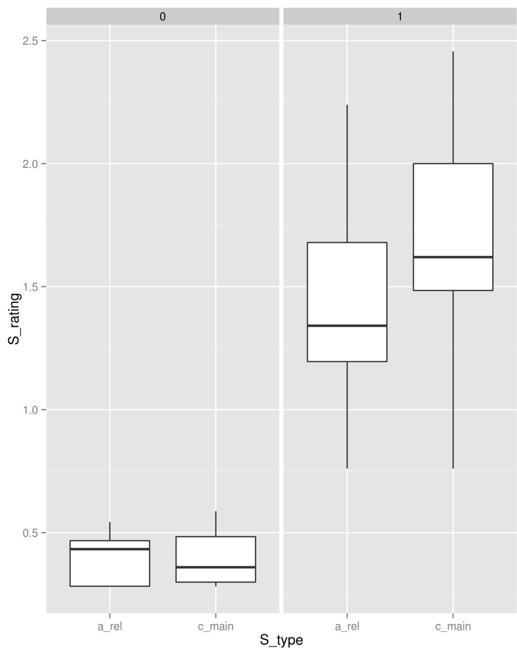

Figure 1: Box plots: absolute sentence rating by sentence type for non-polar (left) and polar (right) adjectives

lexicons, the average scores are identical and the two points lie atop each other.

If we compare the absolute values of individual judges' scores for main and relative clause instances with polar adjectives, we find the pattern shown in Table 2. The results for non-polar adjectives are shown in Table 3.

	0	1	2	3
0	36	45	28	5
1	23	259	191	30
2	3	48	252	59
3	1	4	14	102

Table 2: Confusion matrix for main (columns) and relative clauses (rows) with polar adjectives

For polar adjectives, if a judge does not rate the main and relative clause instances the same, they are 3 times as likely to rate the main clause instance as the more intense type than the relative clause instance. Compare the sum of the cells above the diagonal in Table 2 to the sum of the cells below the diagonal. A chi-square test performed on the confusion matrix in Table 2 is highly significant (X-squared = 733.9092, df = 9, p-value $< 2.2e-16$).

For non-polar adjectives, the likelihood that main and relative clause instances will be rated the same is highest, too. However, if the two are not rated the same, it seems the relative order is more or less random: in half the cases, the relative clause instance was rated as more strongly subjective, in the other half the main clause instance. A chi-square test on the groupings displayed in Table 4 shows that the polar adjectives and the non-polar ones differ significantly (X-squared = 62.1251, df = 2, p-value = 3.234e-14).

	0	1	2	3
0	249	40	8	1
1	48	55	10	0
2	7	8	9	1
3	0	0	3	1

Table 3: Confusion matrix for main (columns) and relative clauses (rows) with non-polar adjectives

	polar	non-polar	Total
M>R	358	60	418
M=R	649	314	963
R<M	93	66	159
Total	1100	440	1540

Table 4: Relative magnitude of main (M) vs. relative (R) clauses with the same adjective, both for polar and non-polar adjectives

To study the relative influence of sentence type and adjective intensity, we fit a cumulative link mixed model to the data (Agresti, 2002). We use clmm2 from R's ordinal package for this purpose (Christensen, 2011).

Our dependent variable is the absolute sentence subjectivity rating. We have two independent variables. The first of these is sentence type, that is, relative clauses versus main clause. The second is adjective intensity. Both these variables are treated as ordinal data. Sentence type has two levels, with class 0 corresponding to relative clauses and class 1 to main clauses, where we expect greater salience of predicates. Adjective intensity is treated as an ordinal variable with four levels by assigning class 0 to adjectives with scores in the range from 0-25, class 1 to adjectives with scores in the range from 26-50, etc.[5] We assume the rater effects are independent and identically distributed random variables.

For the maximum likelihood estimates of the parameters we use the adaptive Gauss-Hermite quadrature method to compute the likelihood function. We

[5]We get very similar results even if we change the class boundaries.

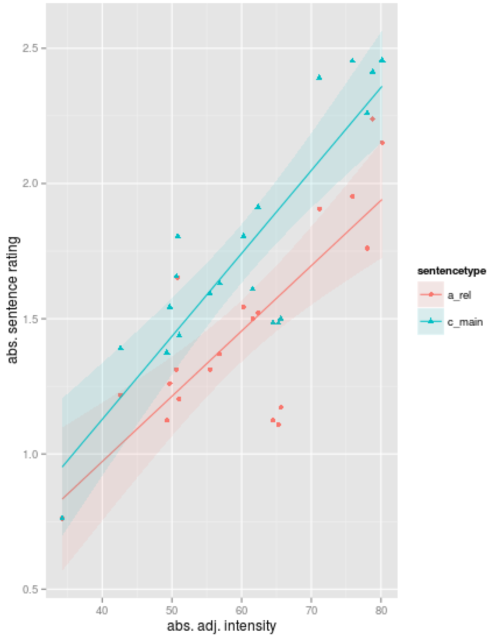

Figure 2: Average absolute sentence rating by average absolute adjective intensity in relative (red dots) and main (green triangles) clauses

use the default setting of 10 quadrature nodes. Significance of model terms was assessed using likelihood ratio tests (a = 0.05), and models were compared with the Akaike Information Criterion (AIC). The condition number of the Hessian was used to assess model fit.

High condition numbers correspond to less well defined models that could be simplified, or models where possibly some parameters are not identifiable. As can be seen from Figure (3), in our case the condition number of the Hessian (61.46202) does not indicate a problem with the model. Figure (3) shows that the coefficients for sentence type (0.5967) and adjective intensity (1.7088) are positive. This indicates that both a sentence type with greater predicate salience and greater intrinsic adjective intensity make a higher sentence subjectivty rating more likely. Additional likelihood ratio tests using the anova method show that sentence type and adjective intensity class are significant terms. The same is true of the variance parameter, rater.

4.3 Discussion

To sum up the analysis so far, we have seen that for our stimuli it seems to be the case that absolute sentence subjectivity ratings depend on sentence type, not only on adjective intensity. Given that the sentences we experiment with contain the same lexical material as well as the same structure, we may conclude that it is the positioning of the subjective expressions in

```
clmm2(location = abssentencerating ~
sentencetype + adjintclass,
random = rater, data = nudat,
Hess = TRUE, nAGQ = 10)

Random effects:
              Var      Std.Dev
rater 0.716407 0.8464083

Location coefficients:
              Estimate Std. Error z value Pr(>|z|)
sentencetype  0.5967   0.0723     8.2504 < 2.22e-16
adjintclass   1.7088   0.0471    36.2611 < 2.22e-16

No scale coefficients

Threshold coefficients:
    Estimate Std. Error z value
0|1  0.9486   0.1123    8.4468
1|2  3.5101   0.1319   26.6095
2|3  6.0575   0.1606   37.7232

log-likelihood: -3071.166
AIC: 6154.332
Condition number of Hessian: 61.46202
```

Figure 3: Fitting a cumulative link mixed model to the data

either the focal main clause or the presupposed relative clause that is responsible for the observed differences. This conclusion is strengthened by the fact that we used multiple pairs of adjectives from different semantic fields and matched the adjectives for frequency.

Through our data elicitation, we found that some of the information in the polarity and intensity dictionaries we used to select our adjectives did not match with the results of our elicitation. For instance, for the polar adjectives, Spearman's rank correlation between the elicited ratings and the information in SentiWS was 0.6782 (cf. Table 1). In addition, 3 of the adjectives that we expected to be objective based on the polarity lexicons behaved close to polar adjectives: *objektiv* 'objective', *wissenschaftlich* 'scientific' and *sachlich* 'matter-of-fact'. In part, this may be due to the background of our subjects: as college students they are taught to value objectivity over subjectivity. Conversely, the item *ungeschickt* 'clumsy' seems not to have been perceived as polar by most subjects, based on the evidence of the sentence subjectivity ratings, even though it is quite strongly polar for SentiWS and moderately so in our own elicitations. The lesson we take away from this is that in future experiments, we should first collect the intensity ratings ourselves before we try to construct pairs with specified differences in intensity.

5 Corpus study

In the rating survey we elicited subjects' *perception* of the intensity of sentences differing only in terms of information structure. We interpreted the results as showing that sentence type influences perceived intensity. However, since the experimental situation is an artificial one–with constructed stimuli and a lack of context–we are interested in complementary evidence that would show that people *use* subjective and objective adjectives in a way that reflects the perceptions we elicited. Accordingly, we performed a corpus study to test the following hypothesis: because main-clause use ensures greater salience of the expression, if speakers want to express opinions with subjective adjectives, they will use them as main clause predicates more often than they would objective adjectives, which do not (directly) serve to express opinions.

We randomly selected 9 adjectives from our pool, 3 each of the negative, positive and objective sets. For each adjective we collected 100 randomly chosen predicative uses in finite clauses from a corpus of German Amazon product reviews (Prettenhofer and Stein, 2010) and classified them as to clause type. Note that we extracted only corpus instances whose word form matched an adjective's invariant predicative form in the positive degree. That is, for e.g. *dumm* 'dumb' we only looked for the word form *dumm* but not for *dümmer*. We extracted the instances from a corpus of reviews so that we could assume that the adjectives are used in a context where the authors generally intend to convey opinions.

We performed the classification manually so as to avoid errors due to erroneous POS-tagging or parsing. In Table 5 we present the results.[6]

	main	relative	other
dumm	73	8	19
entsetzlich	56	1	1
unfreundlich	27	5	9
hübsch	69	4	27
spannend	88	1	11
grandios	97	2	1
wissenschaftlich	79	7	14
geheim	50	21	29
geläufig	59	30	11

Table 5: Main and relative clause occurrences per 100 predicative uses

We can aggregate the numbers for the positive, negative and objective adjectives, as shown in Table 6, and perform a χ^2-test on it. The difference in the distribution of the different types of adjectives across the clause types is highly significant (X-squared = 58.7103, df = 4, p-value = 5.413e-12). As the expected numbers in parentheses show, there are too few instances of relative clause use for the negative and

[6]The "other" category includes e.g. uses in complement clauses, subordinate clauses, etc. Note that *entsetzlich* appalling' and *unfreundlich* 'unfriendly' have fewer than 100 predicative uses in the data we use.

	main	relative	other
negative	156 (149)	14 (20)	29 (30)
positive	254 (225)	7 (30)	39 (45)
objective	188 (225)	58 (30)	54 (45)

Table 6: Aggregate results (expected numbers shown in parentheses)

positive adjectives in the observed sample, while there are too many relative clause uses for the objective adjectives. With respect to main clause uses, objective adjectives do not have enough of them, while especially positive adjectives have more of them than expected. This is also the case for negative adjectives, but less so.

Thus, although the set of adjectives analyzed is small, the results generally support the original hypothesis: for subjective adjectives, placement in main versus relative clauses matters much more than for objective adjectives. In line with that, subjective adjectives are used in relative clauses much less often than objective adjectives.

As shown by the counts in Table 5, *wissenschaftlich* 'scientific' behaves exceptionally. For the polarity dictionaries, this is an objective adjective. However, in our human data elicitation we found that it behaves much like a polar adjective. And this is also what we find here, as can be seen in example (8) from our data:

(8) Dennoch sind die Beispiele und Erklärung esoterisch und nicht **wissenschaftlich**.
'Nonetheless, the examples and the explanation are esoteric and not scientific'.

If we had treated *wissenschaftlich* as a non-objective, positive adjective (as indicated by the dashed line in Table 5), the results of the χ^2-test would have come out even more extreme than they have. However, looking at the corpus data shows that it is not clear that *wissenschaftlich* is inherently positive or negative. Besides frequent uses such as (8), one finds others where *wissenschaftlich* is used negatively.

(9) Das Buch ist natürlich recht **wissenschaftlich** und daher dann und wann vielleicht etwas trocken .
'However, the book is quite scientific and therefore maybe a bit dry every now and then.'

Given the existence of both uses like (8) and (9), it seems correct to say that *wissenschaftlich* is inherently an objective adjective. However, when it is used in context to convey or imply evaluation, it behaves distributionally like an inherently subjective adjective, occurring more often in main clauses than expected and less often in relative clauses.

6 Conclusion

In this paper, we presented the first study showing that in addition to degree modification, negation and modality, information structure has an influence of its own on the salience of subjective expressions. We probed this influence through an online experiment in which we had subjects rate controlled stimuli differing only in information structure. In addition, we performed a corpus study whose results indicate that the differing salience of subjective expressions that was found in the rating experiment also guides people's production of subjective expressions, at least in a context that is geared towards the expression of opinions.

In future work, we plan to extend our study to the occurrence of predicative adjectives in complement clauses as well as to the occurrences of attributive adjectives. Also, since information structure has not been previously identified as a separate variable impacting the perception of subjective expressions, we will want to study in a controlled way how it interacts with other well known variables such as degree modification. Finally, we want to follow up on the question pursued by (Kabadjov et al., 2011) and investigate whether differences in the salience with which an opinion is expressed influence how helpful these opinions are in opinion summarization.

Acknowledgments

We would like to thank our colleague Berry Claus for discussion and advice on the construction of our stimuli and the statistical analysis.

References

Alan Agresti. 2002. *Categorical data analysis*, volume 359. John Wiley & Sons.

Farah Benamara, Baptiste Chardon, Yannick Mathieu, Vladimir Popescu, and Nicholas Asher. 2012. How do negation and modality impact on opinions? In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, pages 10–18. Association for Computational Linguistics.

Heike Bieler, Stefanie Dipper, and Manfred Stede. 2007. Identifying formal and functional zones in film reviews. *Proceedings of the 8th SIGDIAL*, pages 75–78.

Rebecca F Bruce and Janyce M Wiebe. 1999. Recognizing subjectivity: a case study in manual tagging. *Natural Language Engineering*, 5(2):187–205.

Rune Haubo B Christensen. 2011. A Tutorial on fitting Cumulative Link Mixed Models with clmm2 from the ordinal Package. `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.224.8041&rep=rep1&type=pdf`.

Mick P Couper, Michael W Traugott, and Mark J Lamias. 2001. Web survey design and administration. *Public opinion quarterly*, 65(2):230–253.

Gerard de Melo and Mohit Bansal. 2013. Good, great, excellent: Global inference of semantic intensities. *Transactions of the ACL.*

Stefania Degaetano-Ortlieb, Elke Teich, and Ekaterina Lapshinova-Koltunski. 2012. Domain-specific variation of sentiment expressions: a methodology of analysis for academic writing. In *Proceedings of the 12th conference on natural language processing (KONVENS)*, pages 291–295.

Stephan Greene and Philip Resnik. 2009. More than words: Syntactic packaging and implicit sentiment. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 503–511. Association for Computational Linguistics.

Vasileios Hatzivassiloglou and Kathleen R McKeown. 1997. Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, pages 174–181. Association for Computational Linguistics.

Bas Heerschop, Frank Goossen, Alexander Hogenboom, Flavius Frasincar, Uzay Kaymak, and Franciska de Jong. 2011. Polarity analysis of texts using discourse structure. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pages 1061–1070. ACM.

Julian Heister, Kay-Michael Würzner, Johannes Bubenzer, Edmund Pohl, Thomas Hanneforth, Alexander Geyken, and Reinhold Kliegl. 2011. dlexDB–eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 62(1):10–20.

Mijail Kabadjov, Alexandra Balahur, and Ester Boldrini. 2011. Sentiment intensity: Is it a good summary indicator? In Zygmunt Vetulani, editor, *Human Language Technology. Challenges for Computer Science and Linguistics*, volume 6562 of *Lecture Notes in Computer Science*, pages 203–212. Springer Berlin Heidelberg.

Knud Lambrecht. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.

Karo Moilanen and Stephen Pulman. 2007. Sentiment composition. In *Proceedings of the Recent Advances in Natural Language Processing International Conference*, pages 378–382.

Subhabrata Mukherjee and Pushpak Bhattacharyya. 2012. Sentiment analysis in Twitter with lightweight discourse analysis. In *Proceedings of COLING 2012*, pages 1847–1864, Mumbai, India, December. The COLING 2012 Organizing Committee.

Peter Prettenhofer and Benno Stein. 2010. Cross-language text classification using structural correspondence learning. In *Proceedings of the ACL.*

Robert Remus, Uwe Quasthoff, and Gerhard Heyer. 2010. SentiWS - A Publicly Available German-language Resource for Sentiment Analysis. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May. European Language Resources Association (ELRA).

Josef Ruppenhofer, Michael Wiegand, and Jasper Brandes. 2014. Comparing methods for deriving intensity scores for adjectives. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, volume 2: Short Papers*, pages 117–122, Gothenburg, Sweden, April. Association for Computational Linguistics.

Vera Sheinman and Takenobu Tokunaga. 2009. Adjscales: Visualizing differences between adjectives for language learners. *IEICE TRANSACTIONS on Information and Systems*, 92(8):1542–1550.

Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. 2011. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307.

Ulli Waltinger. 2010. GermanPolarityClues: A Lexical Resource for German Sentiment Analysis. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, may. European Language Resources Association (ELRA).

Fei Wang, Yunfang Wu, and Likun Qiu. 2012. Exploiting discourse relations for sentiment analysis. In *Proceedings of COLING 2012: Posters*, pages 1311–1320, Mumbai, India, December. The COLING 2012 Organizing Committee.

Janyce Wiebe and Ellen Riloff. 2005. Creating subjective and objective sentence classifiers from unannotated texts. In *Proceedings of the 6th International Conference on Computational Linguistics and Intelligent Text Processing*, CICLing'05, pages 486–497, Berlin, Heidelberg. Springer-Verlag.

Janyce Wiebe. 2000. Learning subjective adjectives from corpora. In *AAAI/IAAI*, pages 735–740.

Michael Wiegand, Alexandra Balahur, Benjamin Roth, Dietrich Klakow, and Andrés Montoyo. 2010. A survey on the role of negation in sentiment analysis. In *Proceedings of the workshop on negation and speculation in natural language processing*, pages 60–68. Association for Computational Linguistics.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pages 347–354. Association for Computational Linguistics.